- 0.76388, 0.32497, 0.60156
- (+): 1 (0.875), 2 (0.76388), 3 (1), 5 (1), 6(0.6015625) (-): 4 (0.32407), 7 (0.162035), 8 (0), 9 (0.4), 10 (0)

(i)

$$
\begin{aligned}
b_1(x_1) &= \frac{1}{Z}\phi_1(x_1)m_{12}(x_1) \\
&= \frac{1}{Z}\phi_1(x_1)\sum_{x_2}\phi_2(x_2)\psi_{12}(x_1,x_2)\prod_{i\in N_2} m_{2i}(x_2) \\
&= \frac{1}{Z}\phi_1(x_1)\sum_{x_2}\phi_2(x_2)\psi_{12}(x_1,x_2)\sum_{x_3}\phi_3(x_3)\psi_{23}(x_2,x_3)\sum_{x_4}\phi_4(x_4)\psi_{24}(x_2,x_4)
\end{aligned}
$$

(ii)

$$
\begin{aligned}
p(x_1|y_1,y_2,y_3,y_4) &= p(x_1|y_1,x2)\sum_{x_2}p(x_2|y_2,x_3,x_4)\sum_{x_3}p(x_3|y_3)\sum_{x_4}p(x_4|y_4) \\
&= \frac{1}{Z_1}\phi_1(x_1)\frac{1}{Z_2}\sum_{x_2}\phi_2(x_2)\psi_{12}(x_1,x_2)\frac{1}{Z_3}\sum_{x_3}\phi_3(x_3)\psi_{23}(x_2,x_3)\frac{1}{Z_4}\sum_{x_4}\phi_4(x_4)\psi_{24}(x_2,x_4) \\
&= \frac{1}{Z}\phi_1(x_1)\sum_{x_2}\phi_2(x_2)\psi_{12}(x_1,x_2)\sum_{x_3}\phi_3(x_3)\psi_{23}(x_2,x_3)\sum_{x_4}\phi_4(x_4)\psi_{24}(x_2,x_4)
\end{aligned}
$$

(iii) $b_1 = \frac{1}{Z}\begin{bmatrix}1.418791\\1.30009\end{bmatrix}$

$b_2 = \frac{1}{Z}\begin{bmatrix}2.4871\\0.231781\end{bmatrix}$

$b_3 = \frac{1}{Z}\begin{bmatrix}0.418\\2.300881\end{bmatrix}$

$b_4 = \frac{1}{Z}\begin{bmatrix}0.231781\\2.4871\end{bmatrix}$

$b_5 = \frac{1}{Z}\begin{bmatrix}2.12971\\0.589171\end{bmatrix}$

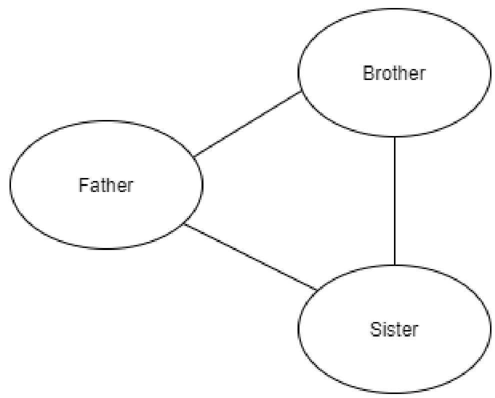
b_2 and b_4 are most likely to agree with their observations which matches with our results. b_5 strongly disagrees with b_3, which in turns disagrees with b_2, correspond to their relations. b_3 and b_4 relation is dominated by the relation between b_3 and b_2. b_1 is only slightly in favor of 1 result, which is the one that agrees with b_2.

One way to achieve 0 loss is to set every vectors to 0, with 1 index being 1 for every e. These embeddings would give no information.

Again, setting every thing to 0, with 1 index being 1 for every e, will achieve minimum loss.

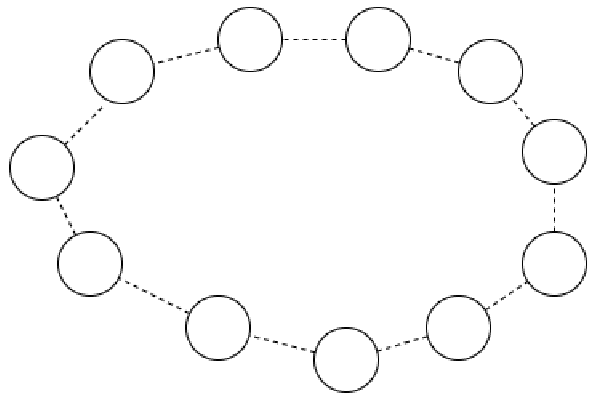
The algorithm can trivially solve L by increasing the norms of the embeddings.

The embeddings for the graph fails when there is a 3 way dependency between e, l , t. For example:



There are no meaningful embeddings in this situation since brother and sister both have the same relation to father so they both equate to $e_{father} + l_{paternity}$ thus $l_{siblings}$ must be a 0 vector.

1. We would need a network of at least 3 layers for the 2 nodes to have different representations



2. Every nodes in this figure can be classified as True as this graph is a cycle of length 10. We can see clearly from this graph that, starting from a random node, we need a message passing chain of 5 to properly cover the cycle to make a decision.

1. $D^{-1}A$
2. $\frac{1}{2}(D^{-1}A + I)$

From 3.2, we found that the transition matrix for the message passing function is $D^{-1}A$. After l layers, the embedding matrix becomes: $(D^{-1}A)^l h$.

We can see that one eigenvector of the transition matrix is $\mathbf{1}$ with corresponding eigenvalue 1. However, 1 is the largest eigenvalue of the matrix (similar proof to `http://web.stanford.edu/class/cs224w/slides/05-spectral.pdf`, slide 11, winter 2019). Therefore, $(D^{-1}A)^l$ converges and thus the embedding converges if l approach ∞.

1. At every step, we examine the nodes in the graph. If at least one of the neighbor of the node has been visited, the note is marked (in this case, have an embedding of 1)
2. The message passing mechanism: $h_i^{l+1} = \mathbf{1}_{(\sum_{j \in N_i} h_j > 0)}(h_i)$

ENZYMES: 120 graphs

CORA: 1000 nodes

Code in the repo. Nothing in this assignment makes sense lol

Results are in Tensorboard dir=./log

Information sheet
CS224W: Machine Learning with Graphs

Assignment Submission Fill in and include this information sheet with each of your assignments. This page should be the last page of your submission. Assignments are due at 11:59pm and are always due on a Thursday. All students (SCPD and non-SCPD) must submit their homework via GradeScope (http://www.gradescope.com). Students can typeset or scan their homework. Make sure that you answer each (sub-)question on a separate page. That is, one answer per page regardless of the answer length. Students also need to upload their code on Gradescope. Put all the code for a single question into a single file and upload it.

Late Homework Policy Each student will have a total of *two* late periods. *Homework are due on Thursdays at 11:59pm PT and one late period expires on the following Monday at 11:59pm PT*. Only one late period may be used for an assignment. Any homework received after 11:59pm PT on the Monday following the homework due date will receive no credit. Once these late periods are exhausted, any assignments turned in late will receive no credit.

Honor Code We strongly encourage students to form study groups. Students may discuss and work on homework problems in groups. However, each student must write down their solutions independently, i.e., each student must understand the solution well enough in order to reconstruct it by him/herself. Students should clearly mention the names of all the other students who were part of their discussion group. Using code or solutions obtained from the web (GitHub/Google/previous year's solutions etc.) is considered an honor code violation. We check all the submissions for plagiarism. We take the honor code very seriously and expect students to do the same.

Your name: ______________________________

Email: ______________________ **SUID:** ____________________

Discussion Group: ______________________________

I acknowledge and accept the Honor Code.

(Signed) ______________________________